DOCUMENT RESUME

ED 252 931 EA 017 418

AUTHOR Tadlock, James; Nesbit, Lamar
TITLE The Relationship of Teacher Evaluation Scores Generated by a Process-Product Evaluation Instrument to Selected Variables.
PUB DATE Nov 84
NOTE 14p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (New Orleans, LA, November 14-16, 1984).
PUB TYPE Reports - Research/Technical (143) -- Statistical Data (110)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Analysis of Covariance; Elementary Secondary Education; Employment Practices; Evaluation Methods; *Evaluation Utilization; Predictor Variables; *Reduction in Force; Regression (Statistics); *School Statistics; *Statistical Distributions; Tables (Data); *Teacher Evaluation
IDENTIFIERS *Jackson Municipal Separate School District MS

ABSTRACT

The Jackson Municipal Separate School District, Mississippi, has instituted a mixed-criteria reduction-in-force procedure emphasizing classroom performance to a greater degree than seniority, certification, and staff development participation. The district evaluation process--measuring classroom teaching performance--generated data for the present study, which is part of a comprehensive 3-year study carried out by the district. This paper examines (1) the distribution of average teacher evaluation scores for elementary, junior, and senior high schools and for the district at large; (2) the independent and composite relationship between evaluation scores and certificate level, staff development points earned, years of teaching experience, and ethnic identity for 1983-84, and the relationship between a reduction-in-force score and the same variables without an evaluation score; and (3) consistency of relationship between criteria and dependent variables between 1982-83 and 1983-84. Findings reveal (1) possible inconsistencies between schools in the evaluation; (2) lack of uniform understanding by evaluators of the "standard score" concept; (3) almost identical results for both years of the study; (4) independence of evaluation score from other known variables; (5) a strong influence of evaluation scores on reduction-in-force value; (6) weak relation between years of experience and evaluation score; (7) no relation between certification level and evaluation score; and (8) weak relationships between staff development points or race and evaluation scores or RIF scores. Recommendations follow, and tables and references are included. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED252931

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

James Tadlock

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE RELATIONSHIP OF TEACHER EVALUATION SCORES GENERATED BY A PROCESS-PRODUCT EVALUATION INSTRUMENT TO SELECTED VARIABLES

by

James Tadlock, Ph. D.
Director of Program Evaluation

and

Lamar Nesbit, Ed. D.
Assistant Superintendent for Personnel Services

JACKSON MUNICIPAL SEPARATE SCHOOL DISTRICT
Jackson, Mississippi
Dr. Robert N. Fortenberry, Superintendent

Mid-South Educational Research Association

November 14-16, 1984
New Orleans, Louisiana

EM017418

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND 1

STATEMENT OF THE PROBLEM AND RELATED QUESTIONS 1

DEFINITIONS 2

PRESENTATION OF DATA AND FINDINGS 3

CONCLUSIONS 5

RECOMMENDATIONS 6

TABLE

I - AVERAGE EVALUATION SCORES RANK ORDERED BY SCHOOL AND BY DIVISION 7

II - INTERCORRELATION MATRIX 9

III - SELECTED CORRELATION COEFFICIENTS FOR 1982-83 AND FOR 1983-84 10

IV - REGRESSION MODELS: CRITERION = EVALUATION SCORE 10

V - REGRESSION MODELS: CRITERION = RIF1 SCORE 10

REFERENCES 11

INTRODUCTION AND BACKGROUND

For several years the Jackson Municipal Separate School District has found it necessary to reduce staff based on pupil enrollment decline, instability of federal funding, and shifting patterns in course selection by secondary students. Prior to 1981, seniority was the primary determinant of reduction-in-force. Since the 1981-82 school year a mixed-criteria reduction-in-force procedure has been used which places primary emphasis on classroom performance with lesser emphasis on seniority, certification, and staff development participation.(1) The district evaluation process was revised prior to the 1982-83 session and has been in use without change since that time. This process establishes a measure of classroom teaching performance and provides data for the purposes of this study. A considerable amount of data has been generated since 1982 and can be utilized to answer certain questions attendant to the RIF process. The strengths, weaknesses, and administrative details of the process are fully documented elsewhere and will not be repeated here.(2) Rather, certain data will be examined in light of specific concerns. This paper will address only a small part of a comprehensive study carried out by the school district over a three-year period.(3)

STATEMENT OF THE PROBLEM AND RELATED QUESTIONS

The following questions were used to guide inquiry and discussion.

1. What is the distribution of average evaluation scores as measured by the Jackson Public Schools teacher evaluation instrument for elementary schools, junior high schools, and senior high schools and for the district for the 1983-84 school year?

2. What is the independent and composite relationship between the criterion of evaluation score and the dependent variables of certificate level, staff development points earned, years teaching experience, and race for 1983-84; and what is the independent and composite relationship

between a reduction-in-force score calculated without an evaluation score and the same independent variables?

3. Is the relationship between criteria and dependent variables consistent between 1982-83 and 1983-84?

DEFINITIONS

Variables are defined and abbreviated as follows:

1. CERTIFICATION LEVEL (CERT).--Each teacher is categorized by the Mississippi State Department of Education certification level attained as A, AA, AAA, OR AAA+. The same certification levels are given respective titles of CERT1, CERT2, CERT3, and CERT4 where group membership designation is necessary.

2. STAFF DEVELOPMENT POINTS (SDP).--The number of staff development points earned by a teacher during the current school year. The current staff development system is highly eclectic and allows the accumulation of "points" for participation in many types of activities.

3. EVALUATION SCORE (SCO).--The derived score calculated from the JPS teacher evaluation instrument. A specific class interval of scores is considered standard with one class interval of scores above standard and two class intervals of scores below standard.

4. RIF1 SCORE (RIF1).--A composite value calculated from the weighted factors of experience out of the district, experience in the district, staff development points earned, degree level, and evaluation score.

5. RIF2 SCORE (RIF2).--A Composite value calculated as RIF1 except evaluation score is not included.

6. RACE (RACE).--Teachers are designated by national racial classifications as minority or non-minority.

7. EXPERIENCE IN DISTRICT (YID).--Actual Teaching experience in this district.

8. EXPERIENCE OUT OF DISTRICT (YOD).--Actual teaching experience out of this district or broken service in this district up to a maximum of 6 years credit.

PRESENTATION OF DATA AND FINDINGS OF THE STUDY

Average evaluation scores for each school and the differences between school and district average evaluation score and the difference between school and division average evaluation score are shown in rank order in Table I. The range of scores for elementary schools is 39.0439 to 46.9762, for junior high schools is 38.2104 to 41.5968, and for senior high schools is 38.7924 TO 42.7613. The range of difference between school average and division average scores for elementary schools is -3.4237 TO 4.5086, for junior high schools is -2.0683 to 1.3181, and for senior high is -1.4277 to 2.5412. Note the spread of scores among elementary schools of 7.93, junior high schools of 3.39, and senior high schools of 3.97. Division average scores of 42.4676 for elementary schools, 40.2787 for junior high schools, and 40.2201 for senior high schools all compare with the district average of 41.2494.

This discussion will give little attention to the distribution of individual teacher evaluation scores. However, the distribution of scores by school are described generally as negatively skewed and leptokurtic. The data are distributed in a wide variety of patterns from school to school within each division. The greatest variety of patterns is found in the elementary division. Distributions of scores for secondary schools are fairly close with two exceptions.

Table II is a matrix of correlations for all variables included in the study for 1983-84. No attempt is made to make a thorough comparison of correlation data between the 1982-83 and 1983-84 school years. However, the correlation of selected variables for both years is compared and reported in table III. The correlations are very close for both years with changes of .02 or less for all variables. The correlations of .82 and .81 between YID and RIF1 is expected since YID is a heavily weighted factor in calculating RIF1. The correlations of .89 and .87 between YID and RIF2 is still higher and can be explained by the fact that RIF2 is not affected by SCO as is RIF1. The correlations of .42 and .44 between SCO and RIF1 is moderate. This can be explained in part with the knowledge that RIF1 is determined only in part by SCO. The correlation of .94 between RIF1 and RIF2 for both years indicates a

high level of agreement, but also indicates that a part of the RIF1 value is being determined by something other than YID, YOD, CERT, and SDP. the low correlations of .14 and .16 between SDP and SCO should also be noted.

Multiple regression models are presented in Tables IV and V with SCO and RIF1 as criteria. The models selected present information to answer questions previously raised. The correlation coefficients shown in Table II form a matrix showing the simple correlation of each variable with every other variable. From the descriptive information in the correlation matrix, multiple regression models were formulated with each criterion and appropriate independent variables. These models can be summarized under two categories: (1) Full model: All predictors are included in this equation. (2) Restricted (Partial) Model: One or more of the predictors included in the full model is (are) omitted from this equation. For each model an iterative process selects the single variable with the highest predictive ability (validity) and continues to select new variables for previously selected variables that will maximally increase predictive ability when combined with the previously included variables(s). A final step calculates the total R (multiple correlation coefficient) and RSQ (predictive efficiency) for the model. Although this report is not concerned with establishing a set of predictive formulae, this technique will establish the strength of relationship between the criterion variable of evaluation score and stated dependent variables. The calculated RSQ can be considered as the proportion of variance in the criterion that can be accounted for by the dependent variable(s). The difference in RSQ between a full model and restricted model is the amount of variance that is "independent" of other variables in the model. This co-variance technique tests the effect of the variable (or variables) omitted from the larger model without the effect of "overlap" from other variables and adds considerable power to interpretation of results. The models shown in Tables IV and V list the variance accounted for by a full model and reduced models with evaluation score as criterion in Table IV and RIF1 score as criterion in Table V. The composite amount of variance in SCO

that is accounted for by CERT, SDP, YID and YOD IS .0433. The independent amount of variance accounted for by dependent variables is shown in Table IV. Obviously, these variables contribute very little to determining evaluation scores.

The composite amount of variance in RIF1 that is accounted for by SDP, CERT, YID and YOD, and SCO IS .9942. The independent amount of variance accounted for by dependent variables is shown in Table V. The iteration sequence for this model merits several observations. The first variable with the highest correlation selected is YID with an RSQ of .6510. This value is obtained from a simple correlation (correlation squared) and indicates the amount of variance in RIF1 that can be determined from YID. The next iteration adds SCO for a combined RSQ OF .8062. This value also indicates the variance in RIF1 that can be accounted for by the addition of SCO to YID. Addition of SCO to the equation raised the predictive validity from .6510 to .8062. the next three iterations add CERT membership vectors for a combined RSQ of .9695. Again, the addition of CERT to SCO and YID raised the amount of variance in RIF1 that can be accounted for from .8062 to .9695. The next two iterations add SDP and YOD to bring the combined RSQ to .9912. The remaining iterations are readjustments in weights from variables entered in previous iterations and brings the final RSQ to .9942.

CONCLUSIONS

1. The variety of distributions of scores between schools and between divisions point out possible inconsistencies between schools in the evaluation process. Some variation is to be expected. The degree of variation should cause concern, particularly between divisions.

2. The leptokurtic negatively skewed distributions of scores for schools, divisions, and the district suggests the lack of uniform understanding by evaluators of the "standard score" concept.

3. The correlation data presented here for both years indicates almost identical results. The strength of relationships identified for 1982-83 are verified with this replication.

4. The evaluation score as calculated is a unique value and cannot be determined through knowledge of the other variables included in the study. This conclusion does not necessarily validate the evaluation instrument for the purpose intended but it does establish beyond question that the evaluation score is independent of other known variables.

5. The evaluation score continues to make considerable contribution toward calculation of the RIF1 value. The independent contribution of .1073 may seem to be rather low, but this value must be considered in conjunction with the composite contribution of all variables. The amount and the power of this value seems to be highly significant from a practical point of view.

6. Years experience teaching still contributes heavily to determination of the RIF1 value but the relationship between years experience and evaluation is very weak.

7. Certification level contributes minimally to the RIF1 value but does not affect the evaluation score.

8. The variables of SDP and RACE have extremely weak relationships with SCO or RIF1.

RECOMMENDATIONS

1. Study of the evaluation process should be continued. The progress made in instrument design and evaluator training is recognized. But, until consistency of evaluation between schools (evaluators) is a reality, the process is subject to criticism for which there is limited defense.

2. Although race is not a factor in determining evaluation scores for the district as a whole, individual school data should be examined to determine if race has affected evaluation scores.

3. Changes in the district staff development program are recommended. Other values may accrue to the individual or the district from the current process, but the accumulation of staff development points make an insignificant contribution to evaluation scores and RIF scores.

TABLE I

AVERAGE EVALUATION SCORES FOR SCHOOLS
RANK ORDERED BY SCHOOL/DIVISION DIFFERENCE

SCHOOL	NUMBER OF TEACHERS	AVERAGE EVALUATION SCORE	SCHOOL/ DISTRICT DIFFERENCE	SCHOOL/ DIVISION DIFFERENCE
ELEM-1	13	46.9762	5.7268	4.5086
ELEM-2	24	46.2526	5.0032	3.7850
ELEM-3	16	45.3196	4.0702	2.8520
ELEM-4	10	44.9191	3.6697	2.4515
ELEM-5	17	44.8189	3.5695	2.3513
ELEM-6	15	44.3576	3.1082	1.8900
ELEM-7	16	44.2859	3.0365	1.8183
ELEM-8	28	44.1250	2.8756	1.6574
ELEM-9	27	44.0407	2.7913	1.5731
ELEM-10	21	44.0128	2.7634	1.5452
ELEM-11	15	43.8419	2.5925	1.3743
ELEM-12	17	43.5769	2.3275	1.1093
ELEM-13	15	43.1768	1.9274	0.7092
ELEM-14	27	42.8057	1.5563	0.3381
ELEM-15	35	42.7930	1.5436	0.3254
ELEM-16	19	42.7550	1.5056	0.2874
ELEM-17	12	42.6474	1.3980	0.1798
ELEM-18	24	42.6425	1.3931	0.1749
ELEM-19	20	42.6166	1.3672	0.1490
ELEM-20	14	42.4684	1.2190	-0.0008
ELEM-21	22	42.1433	0.8939	-0.3243
ELEM-22	12	42.1164	0.8670	-0.3512
ELEM-23	19	41.8731	0.6237	-0.5945
ELEM-24	19	41.8170	0.5676	-0.6506
ELEM-25	19	41.7393	0.4899	-0.7283
ELEM-26	28	41.5631	0.3137	-0.9045
ELEM-27	26	41.3012	0.0518	-1.1664
ELEM-28	17	40.9042	-0.3452	-1.5634
ELEM-29	13	40.8845	-0.3649	-1.5831
ELEM-30	18	40.7953	-0.4541	-1.6723
ELEM-31	27	40.5287	-0.7207	-1.9389
ELEM-32	20	40.3410	-0.9084	-2.1266
ELEM-33	21	40.2020	-1.0474	-2.2656
ELEM-34	14	39.9807	-1.2687	-2.4869
ELEM-35	15	39.8102	-1.4392	-2.6574
ELEM-36	16	39.2331	-2.0163	-3.2345
ELEM-37	21	39.0439	-2.2055	-3.4237
ELEM TOTALS	712	42.4676	1.2182	

(CONTINUED OVER)

TABLE I (CONTINUED)

AVERAGE EVALUATION SCORES FOR SCHOOLS
RANK ORDERED BY SCHOOL/DIVISION DIFFERENCE

SCHOOL	NUMBER OF TEACHERS	AVERAGE EVALUATION SCORE	SCHOOL/ DISTRICT DIFFERENCE	SCHOOL/ DIVISION DIFFERENCE
JRHI-1	40	41.5968	0.3474	1.3181
JRHI-2	47	41.5635	0.3141	1.2848
JRHI-3	39	41.4325	0.1813	1.1538
JRHI-4	53	41.2982	0.0488	1.0195
JRHI-5	48	40.3928	-0.8566	0.1141
JRHI-6	21	40.3607	-0.8887	0.0820
JRHI-7	46	40.0418	-1.2076	-0.2369
JRHI-8	43	38.6815	-2.5679	-1.5972
JRHI-9	44	38.6679	-2.5815	-1.6108
JRHI-10	33	38.2104	-3.0390	-2.0683
JRH TOTALS	414	40.2787	-0.9707	
SRHI-1	41	42.7613	1.5119	2.5412
SRHI-2	46	41.7374	0.4880	1.5173
SRHI-3	48	40.7239	-0.5255	0.5038
SRHI-4	57	39.9795	-1.2699	-0.2406
SRHI-5	43	38.9623	-2.2871	-1.2578
SRHI-6	41	38.9329	-2.3165	-1.2872
SRHI-7	54	38.7924	-2.4570	-1.4277
SRH TOTALS	330	40.2201	-1.0293	
ADMINISTRATIVE	65	39.2755	-1.9739	
GRAND TOTAL	1521	41.2494		

TABLE II

INTERCORRELATION MATRIX

VARIABLE	1 SDP	2 CERT1	3 CERT2	4 CERT3	5 CERT4	6 YID
1 SDP	1.0000	0.0269	0.0749	-0.0511	-0.0169	0.0050
2 CERT1	0.0269	1.0000	-0.0699	-0.1381	-0.0949	0.0827
3 CERT2	0.0749	-0.0699	1.0000	-0.4711	-0.3237	0.0362
4 CERT3	-0.0511	-0.1381	-0.4711	1.0000	-0.6397	0.0510
5 CERT4	-0.0169	-0.0949	-0.3237	-0.6397	1.0000	-0.2781
6 YID	0.0050	0.0827	0.2309	0.0510	-0.2781	1.0000
7 YOD	-0.0105	0.0167	0.0362	0.0249	-0.0631	-0.1509
8 SCO	0.1623	0.0471	0.0652	0.0324	-0.1051	0.0558
9 RIF1	0.1949	0.2362	0.4766	0.0822	-0.5672	0.8068
10 RIF2	0.1549	0.2444	0.5037	0.0784	-0.5906	0.8730
11 RACE	0.1038	-0.0139	-0.0113	-0.1299	0.1548	-0.1968
MEANS =	67.0965	0.0201	0.1925	0.4821	0.3053	13.1231
SIGMAS =	25.1134	0.1402	0.3942	0.4997	0.4605	8.1310

VARIABLE	7 YOD	8 SCO	9 RIF1	10 RIF2	11 RACE
1 SDP	-0.0105	0.1623	0.1949	0.1546	0.1038
2 CERT1	0.0167	0.0471	0.2362	0.2444	-0.0139
3 CERT2	0.0362	0.0652	0.4766	0.5037	-0.0113
4 CERT3	0.0249	0.0324	0.0822	0.0784	-0.1299
5 CERT4	-0.0631	-0.1051	-0.5672	-0.5906	0.1548
6 YID	-0.1509	0.0558	0.8068	0.8730	-0.1968
7 YOD	1.0000	-0.0013	0.0240	0.0261	-0.0140
8 SCO	-0.0013	1.0000	0.4384	0.1139	0.0609
9 RIF1	0.0240	0.4384	1.0000	0.9406	-0.1464
10 RIF2	0.0261	0.1139	0.9406	1.0000	-0.1848
11 RACE	-0.0140	0.0609	-0.1464	-0.1848	1.0000
MEANS =	2.8238	41.2008	59.4816	18.2879	0.4545
SIGMAS =	2.4396	2.9603	9.1106	8.2302	0.4979

TABLE III

CORRELATION COEFFICIENTS FOR SELECTED VARIABLES

VARIABLES	CORRELATION COEFFICIENTS 1982-83	1983-84
YID---RIF1	.82	.81
YID---RIF2	.89	.87
RIF1--RIF2	.94	.94
SCO---RIF1	.42	.44
SDP---SCO	.14	.16

TABLE IV

REGRESSION MODELS
CRITERION = EVALUATION SCORE

CRITERION	VARIABLES INCLUDED IN THE MODEL	RSQ	DIFF*
SCO	SDP--CERT--YID,YOD--RACE	.0433	--
SCO	(SDP)-CERT--YID,YOD--RACE	.0208	.0225
SCO	SDP-(CERT)-YID,YOD--RACE	.0326	.0106
SCO	SDP--CERT-(YID,YOD)-RACE	.0422	.0011
SCO	SDP--CERT--YID,YOD-(RACE)	.0390	.0042

TABLE V

REGRESSION MODELS
CRITERION = RIF1 SCORE

CRITERION	VARIABLES INCLUDED IN THE MODEL	RSQ	DIFF*
RIF1	SDP--CERT--YID,YOD--SCO	.9842	--
RIF1	(SDP)-CERT--YID,YOD--SCO	.9818	.0125
RIF1	SDP-(CERT)-YID,YOD--SCO	.8446	.1497
RIF1	SDP--CERT-(YID,YOD)-SCO	.6170	.3773
RIF1	SDP--CERT--YID,YOD-(SCO)	.8869	.1073

*THE DIFF (DIFFERENCE) IN RSQ REPRESENTS THE PROPORTION OF VARIANCE IN THE CRITERION ACCOUNTED FOR BY THE VARIABLE(S) OMITTED FROM THE MODEL. THOSE VARIABLES OMITTED FROM THE FULL MODEL ARE INDICATED AS ENCLOSED IN PARENTHESES.

REFERENCES

1. Jackson Municipal Separate School District. Manual for Teacher Evaluation. Jackson, Mississippi, 1982.

2. Manatt, Richard, and Stow, Shirley. A Matrix of Teacher Performance Areas and Their Criteria Selected by the School Organizations in the School Improvement Model for Use During the 1981-82 School Year. Unpublished paper: Iowa State University, Ames Iowa, 1982.

3. Nesbit, Lamar, and Tadlock, James. Teacher Evaluation and Reduction-in-Force. Unpublished study by the Jackson Municipal Separate School District: Jackson, Mississippi, 1984.

4. Nesbit, Lamar. A Study of the Relationships between Teacher-related and Evaluator-related Variables and Two Measures of Teaching Performance. Unpublished Ed. D. dissertation: University of Mississippi, 1983.